APPLICATION FOR UNITED STATES PATENT

FOR

ADAPTIVE CONFIGURATION OF PLATFORM

Inventor
Alex A. Lopez-Estrada

Attorney Docket No.: 110349-133006
IPG No: P16600

Prepared by Schwabe, Williamson & Wyatt, PC
Pacwest Center, Suites 1600-1900
1211 SW Fifth Avenue
Portland, Oregon 97204.
(503) 222-9981

Express Mail Label No. EL973637359US
Date of Deposit: December 15, 2003

ADAPTIVE CONFIGURATION OF PLATFORM

Technical Field

The present invention is related to the field of data processing, and in particular, to the adaptation of a data processing platform for different uses.

BACKGROUND

Increasingly, a number of the embedded market segments, such as networking, imaging, industrial computers, and interactive clients, has shifted from utilizing special purpose fixed functionality application specific integrated circuits (ASIC) or components, to standard integrated circuits or components, including general-purpose processors, or platforms with general-purpose processors, input/output peripherals and a "basic" operating system (OS).

However, performance of these general-purpose platforms in the various specific embedded market segments remain a significant issue, as it is difficult, if not virtually impossible, to configure a general-purpose platform for optimal performance in multiple embedded market segments.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the present invention will be described by way of the accompanying drawings in which like references denote similar elements, and in which:

Figure 1 illustrates an overview of an embodiment of the present invention;

Figure 2 illustrates a portion of the operational flow of the analyzer of **Fig. 1** in selecting a set of configuration parameter values, if appropriate, to configure the platform of **Fig. 1**, in accordance with one embodiment;

Figure 3 illustrates a portion of the operational flow in determining whether a workload sufficiently resembles a reference workload, in accordance with one embodiment; and

Figure 4 illustrates a computer system suitable for use to practice one or more aspects of an embodiment of the present invention.

DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION

In the following description, various aspects of embodiments of the present invention will be described. However, it will be apparent to those skilled in the art that other embodiments may be practiced with only some or all of the described aspects. For purposes of explanation, specific numbers, materials and configurations are set forth in order to provide a thorough understanding of the embodiments. However, it will be apparent to one skilled in the art that other embodiments may be practiced without the specific details. In other instances, well-known features are omitted or simplified in order not to obscure the description.

Various operations will be described as multiple discrete operations in turn, in a manner that is most helpful in understanding the embodiments, however, the order of description should not be construed as to imply that these operations are necessarily order dependent. In particular, these operations need not be performed in the order of presentation.

The phrase "in one embodiment" is used repeatedly. The phrase generally does not refer to the same embodiment, however, it may. The terms "comprising", "having" and "including" are synonymous, unless the context dictates otherwise.

Figure 1 illustrates an overview of an embodiment of the present invention. As shown, embodiment **100** may include a platform **102** and an analyzer **104** coupled to each other. Platform **102** may include in particular execution resources **110**, workload **112** and monitor **114**, operatively coupled to each other as shown. Analyzer **104**, on the other hand, may include in particular, resemblance analysis function **116** and sets of configuration parameters values **118**.

Execution resources **110** may be employed to execute workload **112**. Execution resources **110** represent a broad range elements employed to form platforms, including but are not limited to processors, in particular, general-purpose processors, volatile and/or non-volatile storage, I/O peripherals, and OS.

Workload **112** may be any workload, including in particular, but not limited to, those workloads that historically had employed embedded systems, such as networking, imaging, industrial computers, interactive clients, and so forth.

Monitor **114** may be employed to monitor one or more performance events associated with execution of workload **112** by platform **102**. The performance events may include events measured by one or more processor, OS and/or chipset counters. Examples of these performance events include, but are not limited to, clockticks, instructions retired, bus accesses, L2 cache misses, load instructions retired, mispredicted branches retired, branches retired, read operations performed, write operations performed, trace cache misses, translation look-aside buffer load, read operation misses, context switches, soft interrupts, and so forth.

Resemblance analysis function **116** of analyzer **104** may be employed to analyze whether workload **112** sufficiently resembles one of one or more reference workloads. As will be described more fully below, in various

embodiments, the determination may be based at least in part on the performance events observed during monitoring of platform **102**'s execution of workload **102**, and corresponding performance events during prior executions of the reference workloads.

The one or more reference workloads may be workloads for which configuration parameter values **118** are pre-selected for configuring platform **102** to execute the corresponding workloads. Examples of reference workloads may include, but are not limited to, one or more of a route look-up workload, a OSPF workload, a JPEG codec workload, a 3DES encryption/decryption workload, an AES encryption/decryption workload, an IP packet forwarding workload, a H.323 speech codec workload, and so forth.

Workloads **112** may be actual or representative workloads. In other words, the earlier described monitoring, analyses, and so forth may be performed for an operational platform **102** or a "test" platform **102**. Representative workloads may be selected based on the target market segment where the "test" platform **102** will be utilized to resemble "typical" market applications. These workloads may be further categorized by the system components or execution resources **110** they exercise. For example, for processor compute bound applications, the representative workloads utilized may consist of low level functions that execute directly within the processor, i.e. from the processor cache, without exercising peripheral components or agents attached to the processor. For system level applications where multiple components are exercised, as in the case of memory or interrupt bound applications, the representative workloads may be selected to exercise such components.

The configuration parameter values may be pre-established with prior executions of the reference workloads. Examples of configuration parameter

values may include, but are not limited to, one or more OS related settings, such as paging size, buffer sizes, memory allocation policies, and so forth, as well as one or more processor related settings, such as whether a second physical processor, logical processor or processing core should be enabled, and chipset related settings, such as arbitration policies. An example approach to pre-determine configuration parameter values will be further described later.

Thus, during operation, execution of workload **112** by platform **102** may be monitored by monitor **114**. In particular, monitor **114** may monitor for one or more performance events. The observed performance events may be provided to analyzer **104** to analyze and determine whether platform **102** may be re-configured to enhance performance.

Still referring to **Fig. 1**, in various embodiments, each of platform **102** and analyzer **104** may include a networking interface (not shown), coupling platform **102** and analyzer **104** to each other, via a local area network. In alternate embodiments, the networking interfaces may couple platform **102** and analyzer **104** to each other, via a wide area network.

Further, analyzer **104**, in various embodiments, may be hosted by a host computing device. Moreover, monitor **114** may be implemented as an integral part of analyzer **104** monitoring platform **102** remotely instead.

On the other hand, in alternate embodiments, analyzer **104** may be an integral part of platform **102**.

In yet other embodiments, as will be described in more detail below, analyzer **104** may be practiced without resemblance analysis function **116**.

Referring now to **Fig. 2**, a portion of the operational flow of analyzer **104** in accordance with one embodiment, is illustrated. For the embodiment, analyzer **104** includes resemblance analysis function **116**. As shown, on receipt of the

performance events from monitor **114**, block **202**, resemblance analysis function **116** may determine whether workload **112** resembles at least one of the one or more reference workloads, block **204**. The determination may be performed based at least in part on the performance events received (i.e. performance events observed during the monitoring), and performance events observed during prior executions of the reference workloads.

If none of the one or more reference workloads is determined to sufficiently resemble workload **112**, block **206**, no selection is made of the configuration parameter value sets, block **208**.

On the other hand, if one of the one or more reference workloads is determined to sufficiently resemble workload **112**, block **206**, the corresponding set of one or more configuration parameter values **118** may be selected, block **210**, and provided **212** to platform **102** to be applied to configure platform **102**.

Figure 3 illustrates a portion of the operational flow of resemblance analysis function (RAF) **116** for determining whether a workload resembles any of the reference workloads, in accordance with one embodiment. As illustrated, RAF **116** first selects one of the reference workloads for analysis, block **302**. Then, RAF **116** determines a correlation metric between the workload and the currently selected reference workload, block **304**.

In various embodiments, RAF **116** may determine the correlation metric as a ratio between the covariance of the performance events observed during execution of the workload, and observed during prior execution of the reference workload, and the product of the standard deviations of the respective performance events observed. Mathematically, the correlation metric may be expressed as follows:

Let X be a vector corresponding to a set of performance events and Y_i be the i^{th} reference workload vector of performance events. The i^{th} correlation coefficient (ρ_i) is given by:

$$\rho_i = \frac{Cov(X, Y_i)}{sX \cdot sY_i}$$

where $Cov(X,Y)$ is the covariance coefficient, and sX and sY_i are the standard deviations of the vectors X and Y_i:

$$Cov(X, Y_i) = \sum_{n=1}^{N} (x[n] - \bar{x}) \cdot (y_i[n] - \bar{y}_i)$$

$$sX = \frac{1}{N} \cdot \sum_{n=1}^{N} (x[n] - \bar{x})$$

$$sY_i = \frac{1}{N} \cdot \sum_{n=1}^{N} (y_i[n] - \bar{y}_i)$$

where N is the number of events in the vector and $\bar{x}$ and $\bar{y}_i$ are the vector means given by:

$$\bar{x} = \frac{1}{N} \sum_{n=1}^{N} x[n]$$

$$\bar{y}_i = \frac{1}{N} \sum_{n=1}^{N} y_i[n]$$

Under this design, the correlation coefficient will fall between –1.0 and 1.0. The closer a correlation coefficient is to 1.0, the more correlated two vectors are, indicating that both data sets vary together.

Continuing to refer to **Fig. 3**, for the embodiment, upon determining the correlation metric between the workload and the currently selected reference workload, RAF **116** determines if more resemblance analysis is to be performed for at least one other reference workload. If so, RAF **116** returns to block **302**, and continues from there as earlier described.

Eventually, RAF **116** would have computed the correlation metrics for all reference workloads.

At such time, RAF **116** determines whether any of the correlation metrics exceeds a correlation threshold, block **308**. If no correlation metric exceeds a correlation threshold, the workload will be considered as having insufficient resemblance to any of the reference workloads, block **310**.

On the other hand, if one of the correlation metrics exceeds a correlation threshold, block **308**, RAF **116** selects the reference workload with the correlation metric greater than the threshold as the resembled workload, block **312**.

Referring back to **Fig. 1**, as alluded to earlier, in various alternate embodiments, analyzer **104** may be practiced without resemblance analysis function **116**. For some of these alternate embodiments, analyzer **104** may be practiced with e.g. a direct lookup function (not shown) instead. The direct lookup function may generate a lookup index based on the performance events observed, and employ the lookup index to lookup (select) one of the one or more sets of pre-established configuration parameter values instead.

The direct lookup function may generate the lookup index by e.g. evaluating an index function in view of the performance events observed. The index function may e.g. be a hashing function. Alternatively, the index function may apply a number of corresponding weights to the performance events observed to generate the index. The corresponding weights may be determined

via a number of quantitative techniques, including but are not limited, neural network techniques, co-factor analysis, and so forth.

Additionally, in various embodiments, configuration parameter values may be determined by selecting a combination of configuration parameter values that yield the lowest processor cycles per unit of work performed by a reference workload. More specifically, the configuration parameter values may be pre-determined via Design of Experiments (DOE) techniques such as full-factorial analysis or fractional factorial analysis. In the former case, all possible combinations of the configuration parameters may be assembled in a matrix, and the performance response (e.g. total number of processor cycles incurred) is measured for each combination. The combination that results in the lowest total processor cycles may be selected as the pre-determined configuration parameter values.

To further illustrate, consider an embodiment with 3 configuration parameters: (1) number of logical processors (one or two), (2) memory page sizes (4KB or 4MB) and (3) hardware pre-fetch mode (enabled or disabled). The total number of combinations is given by $Levels^{Num\ of\ Parms}$. For this example, there are a total of 3 parameters, each with 2 levels; therefore, there are a total of 2^3 or 8 possible combinations. The full factorial matrix is

Combination	Number of Logical Processors	Memory Page Size	Pre-fetcher Mode	Measured Response (total cycles)
1	1	4 KB	Enabled	Y1
2	1	4 KB	Disabled	Y2
3	1	4 MB	Enabled	Y3

4	1	4 MB	Disabled	Y4
5	2	4 KB	Enabled	Y5
6	2	4 KB	Disabled	Y6
7	2	4 MB	Enabled	Y7
8	2	4 MB	Disabled	Y8

In one embodiment, the configuration parameter values that yield the smallest measured response, MIN(Y1, Y2 ..., Y8) are selected as the pre-determined parameter values.

Further, platform **102** may be used for heterogeneous or periodic changed workloads. For example, a set-top box may be used as a DVD player (a video decoding emphasized workload) at one point in time, an audio player (an audio decoding emphasized workload) in another point in time, or web browsing (a TCP/IP and/or encryption/decryption emphasized workload) in yet another point in time, or combinations thereof. Accordingly, the monitoring, analyses, adaptation etc. may be repeated in view of the frequency the workload changes. In other words, the platform may be adapted periodically with a frequency and adaptation pattern that substantially matches the expected change in workload. In alternate embodiments, a weighted approach (based on the expected heterogeneous workload) may be practiced instead.

Figure 4 illustrates a computer system suitable for use to practice one or more aspects of an embodiment of the present invention. As illustrated, computing device **400** may include one or more processors **402**, system memory **404**, mass storage devices **406**, other I/O devices **408** and communication interface **410**, coupled to each other via system bus **412** as shown.

Processor **402** is employed to execute a software implementation of analyzer **104**, and optionally, monitor **114**. Processor **402** may be any one of a number of processors known in the art or to be designed. Examples of suitable processors include but are not limited to microprocessors available from Intel Corp of Santa Clara, CA.

Memory **404** may be employed to store working copies of analyzer **104**, and optionally, monitor **114**. Memory **404** may be Dynamic Random Access Memory (DRAM), Synchronous DRAM (SDRAM) or other memory devices of the like.

Mass storage devices **406** may be employed to persistently store data, including e.g. a persistent copy of analyzer **104**, and optionally, monitor **114**. Examples of mass storage devices **406** include but are not limited to hard disks, CDROM, DVDROM, and so forth.

Other I/O devices **408** may be employed to facilitate other aspects of input/output. Examples of other I/O devices **408** include but are not limited to keypads, cursor control, video display and so forth.

Communication interface **410** may be employed to facilitate e.g. network communication with other devices. For these embodiments, network communication interface **410** may be wired based or wireless. In various embodiments, network communication interface **410** may support one or more of a wide range of networking protocols.

Accordingly, various novel methods and apparatuses for adaptively configuring a platform have been described. While the present invention has been described in terms of the foregoing embodiments, those skilled in the art will recognize that the invention is not limited to the embodiments described. Other embodiments may be practiced with modification and alteration within the

spirit and scope of the appended claims. Accordingly, the description is to be regarded as illustrative instead of restrictive.